The Third-Difference Approach to Modified Allan Variance

Charles A. Greenhall

Jet Propulsion Laboratory

California Institute of Technology

4800 Oak Grove Dr, MS 298-100

Pasadena, California 91109, USA

February 7, 1995

ABSTRACT

This study gives strategies for estimating the modified Allan variance (mvar), and formulas for computing the equivalent degrees of freedom (edf) of the estimators. A third-difference formulation of mvar leads to a tractable formula for edf in the presence of power-law phase noise. The effect of estimation stride on edf is tabulated. First-degree rational-function approximations for edf are derived, and their performances tabulated. A theorem allowing conservative estimates of edf in the presence of compound noise processes is given.

I. INTRODUCTION

Allan variance (AVAR) and modified Allan variance (MVAR) are statistical measures of fractional frequency instability. They are both used extensively to measure and characterize the stability performance of clocks, oscillators, and systems for disseminateing time and frequency [1][10][11][12]. Let us give brief definitions. The raw data for these measures comprise a sequence x_n of time residuals, say from a comparison of two clocks or a phase comparison of two oscillators. We assume here that the samples x_n are evenly spaced in time, with *sample period* τ_0. Let an averaging *time* $\tau = m\tau_0$ be given, where m is an integer. The Allan variance, denoted by $\sigma_y^2(\tau)$, is defined as $1/(2\tau^2)$ times the time average or mathematical expectation of the squares of second differences, with step m, of the sequence x_n. Modified Allan variance, denoted by mod $\sigma_y^2(\tau)$, is defined in the same way, except that the x_n sequence is replaced by the sequence $\bar{x}_n(m)$ of moving averages

$$\bar{x}_n(m) = \frac{1}{m} \sum_{j=0}^{m-1} x_{n-j}. \tag{1}$$

By virtue of the second difference in their definitions, stable statistical estimates of AVAR and MVAR can be accumulated in the presence of a class of of phase noise models, the processes with stationary second increments [10], from which useful fits to the behavior of oscillators, amplifiers, etc., can be selected. Special cases are *power-law* models, associated with spectral densities having the property

$$S_x(f) \sim \text{const}\cdot f^{\beta}$$

as $f \to 0$, where $\beta > -5$. In the usual nomenclature of frequency and time, the noises associated with $\beta = 0, -1, -2, -3, -4$ are called white phase (wh ph), flicker phase (fl ph), white frequency (wh fr), flicker frequency (fl fr), and random-walk frequency (rw fr), respectively. Nonintegral values of β are also allowed; the corresponding noises are called '{fractional",

A log-log plot of measured $\sigma_y(\tau)$ or mod $\sigma_y(\tau)$ vs 7, the familiar σ-τ plot, often indicates phase noise that can be modeled as a linear combination of uncorrelated power-law components, the component associated with β being identified by a straight-line section with slope $\frac{1}{2}(-3-\beta)$. The main advantage of MVAR over AVAR is the increased range of β over which this slope relationship holds: $-5<\beta<-1$ for AVAR, - 5< β < 1 for MVAR [3]. In particular, a mod $\sigma_y(\tau)$ plot can easily distinguish white phase ($\sigma \propto \tau^{-3/2}$) from flicker phase ($\sigma \propto$ ~]). The corresponding asymptotic $\sigma_y(\tau)$ dependencies, 7 "' and $T^{1}\sqrt{\log(a\tau)}$ for some a, can barely be distinguished in practice.

It would seem from (1) that the extra averaging operation that gives MVAR its superior power of discrimination also multiplies the amount of calculation by a factor of m. Previous papers [2] [7], which treat the mechanics of MVAR computation, show how to reduce that factor to 4/3, excluding an initial operation on the data set. The approach given in [7] reformulates the definition of M VAR in terms of third differences of the cumulative sum of the time residuals. Here, after restating this formulation, we apply it to the study of the confidence of estimators" of MVAR in terms of their equivalent degrees of freedom (edf). Tractable expressions for edf in the presence of power-law noise allow extensive numerical trials of estimator parameters, especially the estimation period, the amount by which the estimator summands are shifted in time. The outcome is a practical guideline for estimator design. Simple approximations to the edf of these estimators are constructed and tested, with the aim of providing a convenient package for computing approximate confidence values for most experimental situations. Finally, we show how to obtain conservative confidence values in the presence of phase noise whose spectrum is a sum of power laws.

11. MVAR AND ITS ESTIMATORS

A. Third-Difference Formulation

The definition, calculation, and statistical theory of modified Allan variance are all simplified by an approach that derives MVAR from the cumulative sum of the time residuals x.. We begin wit h the standard formulation. Choose an averaging time $\tau = m\tau_0$, and form the time-residual moving averages $\bar{x}_n(m)$ from (1). Let Δ_m be the backward difference operator, defined by $\Delta_m f_n = f_n - f_{n-m}$ for any sequence f_n. Use the second-difference operator Δ_m^2 to form the MVAR filter output

$$
\begin{aligned}
z_n(m) &= \Delta_m^2 \bar{x}_n(m) \qquad (2)\\
&= \bar{x}_n(m) - 2\bar{x}_{n-m}(m) + \bar{x}_{n-2m}(m)
\end{aligned}
$$

By definition,

$$
\operatorname{mod} \sigma_y^2(\tau) = \frac{1}{2\tau^2} \left\langle z_n^2(m) \right\rangle , \qquad (3)
$$

where ( ) denotes either mathematical expectation E or an infinite time average over n. Note that, although only the variable τ appears, mod $\sigma_y^2(\tau)$ depends on both τ and τ_0. For brevity, we shall occasionally suppress the dependence of $z_n(m)$ on the parameter m.

The third-difference formulation expresses $z_n(m)$ in terms of the sequence w_n defined by

$$
w_0 = 0, \quad w_n = \sum_{j=1}^{n} x_j . \qquad (4)
$$

In terms of w_n, the time-residual averages are given by

$$
\bar{x}_n(m) = \frac{1}{m} \Delta_m w_n = \frac{1}{m} (w_n - w_{n-m}) , \quad n \geq m,
$$

which, combined with (2), gives

$$
z_n(m) = \frac{1}{m} \Delta_m^3 w_n \qquad (5)
$$

$$= \frac{1}{m}\left(w_n - 3w_{n-m} + 3w_{n-2m} - w_{n-3m}\right)$$

for $n \geq 3m$.

Formula (5) has several advantages over (2) for use in (3). The filter taking w_n to $z_n(m)$ has only four taps; the filter taking x_n to $z_n(m)$ has $3m$ taps. The computation of estimates of mod $\sigma_y^2(\tau)$ from third differences of w_n is like the computation of estimates of $\sigma_y^2(\tau)$ from second differences of x_n, and the computation of strided estimates is simplified. Finally, it is easy to construct useful and tractable stochastic models of the w_n sequence. The cost of these advantages is the computation of w_n from the recursion $w_n = w_{n-1} + x_n$.

B. MVAR Estimator with Variable Stride

To estimate MVAR with limited data, the infinite average in (3) is replaced by a finite average of the z_n^2 (m). When computing analogous estimates of AVAR by averaging the squares of $\Delta_m^2 x_n$, it is customary to increase n by either 1 (full overlap) or m (τ overlap). The existing literature on MVAR ([1], for example) usually assumes a step of 1. Here, we allow the step to vary between these extremes. Let us establish some terminology. We specify an ***estimation period*** $\tau_1 = m_1 \tau_0$, where the positive integer m_1 is called the ***estimation stride***, and we consider averages over all available values of $z_{3m+km_1}^2(m)$, $k \geq 0$.

Assume that N time residuals $x_1, x_2, \ldots, x_N$ are available. Then there are $N + 1$ summed values $w_0, w_1, w_2, \ldots, w_N$. Let M be the number of samples of $z_{3m+km_1}(m)$ obtainable from (5). Then M is the largest integer satisfying $3m + (M-1)m_1 \leq N$, namely,

$$M = \left\lfloor \frac{N - 3m + m_1}{m_1} \right\rfloor, \tag{6}$$

where $\lfloor a \rfloor$ denotes the integer part of a. The MVAR estimator to be studied is

$$V = \frac{1}{2\tau^2 M} \sum_{k=0}^{M-1} z_{3m+km_1}^2(m). \tag{7}$$

C. Continuous-Time Analog

A continuous-time analog of this set up yields simple and useful approximations. It is convenient to change the definitions, not only of the underlying noise processes (see below), but also of MVAR and V, by changing discrete-time averages to continuous-time averages. The third-difference approach works here, as well. Let $x(t)$ represent time deviation as a function of time. Write

$$\bar{x}(t;\tau) = \frac{1}{\tau}\int_0^{\tau} x(t-u)\, du,$$

$$z(t;\tau) = \Delta_\tau^2 \bar{x}(t;\tau), \quad w(t) = \int_0^t x(u)\, du.$$

Then

$$\bar{x}(t;\tau) = \frac{1}{\tau}\Delta_\tau w(t),$$

and hence

$$z(t;\tau) = \frac{1}{\tau}\Delta_\tau^3 w(t). \tag{8}$$

Define the continuous-time analog of mod $\sigma_y^2(\tau)$ by

$$\mathrm{mod}^c\,\sigma_y^2(\tau) = \frac{1}{2\tau^2}\left\langle z^2(t;\tau)\right\rangle \tag{9}$$

(identical to Bernier's asymptotic MVAR [3]), and the continuous-averaging estimator V^c by

$$V^c = \frac{1}{2\tau^2 T}\int_0^T z^2(t;\tau)\, dt. \tag{10}$$

Note that if $x(t)$ is available for a duration T_x, then we should let $T = T_x - 3\tau$, the duration of availability of $z(t;\tau)$. Later, to match properties of V^c to those of V, we shall let $T = M\tau_1$, where M is given by (6).

III. Noise Models

The statistical properties of V depend on the random processes chosen to represent the sampled time residuals z_n. Following Walter's treatment of discrete sampling [13], we use an explicit discrete-time power-law model instead of a sampled continuous-time model for our main calculations. This has two advantages. First, we avoid the complications of the interactions among the hardware bandwidth, the sample period, and the averaging time [3] [11]. Second, the discrete-time model works especially well with the third-difference formulation.

Because the measure of estimator confidence to be examined is invariant to scale factors, we use the most convenient scaling for spectral densities to reduce the complexity of constant factors in the generalized autocovariances shown in Table 1. Factors for converting to the standard scaling used by the frequency and time community are given below.

The most critical assumption about the models is the absence of linear frequency drift. We assume that the drift rate either is zero or is known from considerations external to the immediate data set. In the latter case, we can assume that the drift has been removed from the data. In particular, x_n has no long-term quadratic component, w_n has no long-term cubic component, and $z_n(m)$ has mean zero. This assumption will later be repeated at the point where it is needed.

A. Discrete-Time Power Laws

Let the two-sided spectral density of the τ_0-sampled sequence x_n be given by

$$S_x^{\mathrm{d}}(f) = |2\sin(\pi f \tau_0)|^{\beta}, \; |f| < \frac{1}{2\tau_0}. \tag{11}$$

Then $S_x^{\mathrm{d}}(f) \sim |2\pi f \tau_0|^{\beta}$ as $f \to 0$. The process w_n defined by (4) satisfies $\Delta^2 w_n = x_n$. Because the frequency response of the operator Δ^2 is $(2\sin(\pi f \tau_0))^2$, we know that w_n is also a power-law process, with spectral density

$$S_w^{\mathrm{d}}(f) = |2\sin(\pi f \tau_0)|^{\beta-2}.$$

This frequency-domain description of w_n has an equivalent time-domain description, the *generalized autocovariance* (GACV) sequence $R_w^{\mathrm{d}}(n)$, where n runs through all the integers. This concept, whose definition and theory will not be given here, extends the usual notion of autocovariance (ACV) from stationary processes to processes with stationary dth increments [5]. GACVs of continuous-time and discrete-time processes have been used in studies of Allan variance and power-law noise simulation [4] [5] [6] [9]. Table 1 gives formulas for $R_w^{\mathrm{d}}(n)$ for the values of β needed in this study. The formula for nonintegral β in Table 1 is equivalent to the one given by Kasdin and Walter [9] for power laws. Because passage to the limit as β approaches an integer is not straightforward in general, the formulas for integral β are derived from ACVS of stationary power-law processes by solving the difference equation $-\delta_1^2 R(n;\gamma) = R(n;\gamma+2)$ repeatedly, where $R(n;\gamma)$ is the ACV or GACV of a power-law process with exponent γ, and δ_1^2 is the second-order central difference operator with step 1.

Because $\beta > -5$, we know that x_n has stationary second increments, w_n has stationary third increments, and, for each m, $z_n(m)$ is stationary. The ordinary ACV sequence

$$R_z^{\mathrm{d}}(n;m) = \mathrm{E}\ z_{k+n}(m) z_k(m)$$

can be calculated from $R_w^{\mathrm{d}}(n)$ by

$$\begin{aligned} m^2 R_z^{\mathrm{d}}(n;m) &= -\delta_m^6 R_w^{\mathrm{d}}(n) \qquad (12)\\ &= -R_w^{\mathrm{d}}(n-3m) + 6R_w^{\mathrm{d}}(n-2m) - 15R_w^{\mathrm{d}}(n-m) + 20R_w^{\mathrm{d}}(n) \\ &\quad -15R_w^{\mathrm{d}}(n+m) + 6R_w^{\mathrm{d}}(n+2m) - R_w^{\mathrm{d}}(n+3m). \end{aligned}$$

This formula follows from (5) and the theory of GACVs of processes with stationary dth increments.

It is appropriate to note here that (3), Table 1, and (12) give a formula for MVAR in the presence of discrete-time power-law phase noise, namely,

$$\operatorname{mod}\sigma_y^2(\tau) = \frac{1}{2\tau^2}\mathrm{E}\, z_n^2(m) = \frac{1}{2\tau^2}R_z^{\mathrm{d}}(0;m), \tag{13}$$

which, when expanded, is equivalent to a formula of Walter ([13], eq (75)).

The standard power-law scaling used by the frequency and time community is based on a one-sided spectral density, $S_y^+(f) \sim h_\alpha f^\alpha$, of fractional frequency $y = dx/dt$, where $\alpha = \beta + 2$. To convert $R_w^{\mathrm{d}}(n), R_z^{\mathrm{d}}(n;$ m), and mod $\sigma_y^2(\tau)$ to this scaling, multiply them by the factor

$$\frac{h_\alpha}{2}(2\pi)^{-\alpha}\tau_0^{2-\alpha}. \tag{14}$$

B. Continuous-Time Power Laws

Because the continous-time analog given above avoids sampling altogether, continuous-time random-process models are appropriate. Let the two-sided spectral density of $x(t)$ be given by

$$S_x^{\mathrm{c}}(f) = |2\pi f|^\beta, \quad -\infty < f < \infty, \tag{15}$$

with no high-frequency cutoff. Then, since $dw/dt = x$, we know that $w(t)$ is also a power-law process, with spectral density

$$S_w^{\mathrm{c}}(f) = |2\pi f|^{\beta-2}.$$

For $\beta > -5$, the process $w(t)$ has stationary third increments. Its GACV function $R_w^{\mathrm{c}}(t)$ [4] [9] is also given in Table 1. As with the discrete-time model, the process $z(t)$ given by (8) is stationary, with ACV function

$$R_z^{\mathrm{c}}(t;\tau) = \mathrm{E}\, z(u+t;\tau)z(u;\tau)$$

that can be calculated by

$$
\begin{aligned}
\tau^2 R_z^c(t;\tau) &= -\delta_\tau^6 R_w^c(t) \\
&= -R_w^c(t-3\tau) + 6R_w^c(t-2\tau) - 15R_w^c(t-\tau) + 20R_w^c(t) \\
&\qquad -15R_w^c(t+\tau) + 6R_w^c(t+2\tau) - R_w^c(t+3\tau).
\end{aligned}
\tag{16}
$$

A formula for mod' $\sigma_y^2(\tau)$, analogous to (13), is

$$\text{mod}^c\,\sigma_y^2(\tau) = \frac{1}{2\tau^2} R_z^c(0;\tau). \tag{17}$$

Substituting $R_w^c(t)$ from Table 1 into (16), we find from (17) that $\text{mod}^c\,\sigma_y^2(\tau)$ is *exactly* proportional to $\tau^{-3-\beta}$, for – 5<0<1. The same result was derived by Bernier [3] from a frequency-domain integral.

The factor for converting $R_w^c(t), R_z^c(t;\tau)$, and mod' $\sigma_y^2(\tau)$ to standard frequency and time scaling is the same as (14), with *To* set to 1.

IV. Equivalent Degrees of Freedom

By definition, the equivalent degrees of freedom (edf) of a random variable X is defined by

$$\text{edf}\,X = \frac{2\,(\text{E}\,X)^2}{\text{var}\,X}. \tag{18}$$

If X is distributed as a constant multiple of a χ_ν^2 random variable, with ν degrees of freedom, then edf X = ν. For example, the sample variance of n independent, identically distributed Gaussians has n – 1 degrees of freedom. Even if X does not have such a distribution, edf X can still serve as a convenient dimensionless measure of the confidence of X as an estimator of its mean E X. In this study, I take this point of view with regard to *V*, not having investigated the nature of its distribution. Since *V* is the sum of squares of *correlated* mean-zero Gaussians, it is reasonable to assume that *V* is *approximately* distributed as const $\cdot\,\chi^2_{\text{edf}\,V}\cdot$

In this case, approximate confidence intervals for mod $\sigma_y(\tau)$ can be constructed, as described for $\sigma_y(\tau)$ by Howe, All an, and Barnes [8].

A. Discrete Time

Let us compute edf V. By (7) and (13),

$$\mathrm{E}\,V = \frac{1}{2\tau^2} R_z^{\mathrm{d}}(0; m); \tag{19}$$

that is, V *is* unbiased for mod $\sigma_y^2(\tau)$. Also from (7) we have

$$\operatorname{var} V = \frac{1}{(2\tau^2 M)^2} \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} \operatorname{cov}\left(z_{3m+im_1}^2, z_{3m+jm_1}^2\right). \tag{20}$$

To compute the covariances in (20), we assume that $z_n(m)$ is a stationary Gaussian *mean-zero* process. As indicated earlier, the assumption E $z_n(m)$ = O is crucial; in practice, it means that the effect of linear frequency drift on a time scale of order 7- is negligible. Since any two jointly Gaussian mean-zero random variables X and Y satisfy cov (X^2, Y^2) = 2 $(\mathrm{E}XY)^2$, *(20) becomes*

$$\operatorname{var} V = \frac{2}{(2\tau^2 M)^2} \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} \left[R_z^{\mathrm{d}}((i-j)m_1; m)\right]^2. \tag{21}$$

The diagonal $i - j = k$, for $k = 1 - M$ to $M - 1$, contains $M - |k|$ identical term... Summing over these diagonals converts the double sum to a single sum, which, combined with (18) and (19), gives the main formula for edf V:

$$\frac{1}{\operatorname{edf} V} = \frac{1}{M}\left[1 + 2\sum_{k=1}^{M-1}\left(1 - \frac{k}{M}\right)\left(\rho_z^{\mathrm{d}}(km_1; m)\right)^2\right], \tag{22}$$

where

$$\rho_z^{\mathrm{d}}(n; m) = \frac{R_z^{\mathrm{d}}(n; m)}{R_z^{\mathrm{d}}(0; m)}.$$

Formula (22) is analogous to existing formulas for the edf of AVAR estimators (see [6] and references therein). The main difference is that the **ACV** of z is computed from sixth differences of the **GACV** of w instead of fourth differences of the **GACV** of x.

Recall from (12) that each value of $R_z^d(n;m)$ needed in (22) is obtained from seven values of R_w^d (n). If no values of R_w^d (n) are stored in advance, it takes $7M$ evaluations of $R_w^d(n)$ to compute (22). Walter's formula for var V ([1 3], eq.(32)) is a double sum requiring $5(2m-1)(2M-1)$ evaluations of $R_x^d(n)$. In practice, moreover, one can compute and store the values $R_w^d(n), |n| \leq N$, in advance. This shows the advantage of the third-difference approach, which derives mod u:(T) from w_n instead of z.(m).

A note on numerical computation. The ACV $R_z^d(n;m)$ tends to zero as $n \to$ 00, yet is obtained from differences of R_w^d (n), which tends to ∞ with n. Clearly, one should use double precision for evaluating (12). Even so, the computed values of $R_z^d(n;m)$ can deteriorate for large n, especially for nonintegral β, when $R_w^d(n)$ has to be computed from a ratio of Γ functions. I was able to cure this problem by replacing the upper limit $M-1$ of the summation in (22) by K — 1, where K is the smaller of M and $10m/m_1$. (In all actual computations, m/m_1 is assumed to be an integer.)

B. Continuous Time

The computation of edf V^c follows the same pattern. By (10),

$$\mathrm{E}\,V^c = \frac{1}{2\tau^2} R_z^c(0;\tau),$$

and, with the assumption that $z(t;\tau)$, as a function of t, is a stationary Gaussian mean-zero process,

$$\operatorname{var} V^c = \frac{2}{(2\tau^2 T)^2} \int_0^T \int_0^T [R_z^c(t-u;\tau)]^2 \, dt\, du.$$

A change of variables converts the double integral to

$$2\int_0^T (T-t)\,(R_z^c(t;\tau))^2 dt,$$

in which we shall make the further change of variable $t = \tau x$. From Table 1 and (16), it can be verified that

$$\frac{R_z^c(\tau x;\tau)}{R_z^c(0;\tau)} = \frac{R_z^c(x;1)}{R_z^c(0;1)}.$$

(This is a scaling property of continuous-time power-law noise.) Thus, defining

$$\rho_z^c(x) = \frac{R_z^c(x;1)}{R_z^c(0;1)},$$

we obtain

$$\frac{1}{\mathrm{edf}\, V^c} = \frac{2}{p}\int_0^p \left(1 - \frac{x}{p}\right)\left(\rho_z^c(x)\right)^2 dx, \tag{23}$$

where $p = T/\tau$.

V. EFFECT OF ESTIMATION PERIOD

Formula *(22)* was used to generate tables of edf V for combinations of N, m, and m_1. Recall that N is the number of time residuals in the data set, $m = \tau/\tau_0$, where τ is the averaging time, and $m_1 = \tau_1/\tau_0$, the estimation stride, where τ_1 is the estimation period. From here on, we also assume the *divisibility* condition, which says that the estimation period divides evenly into the averaging time, that is,

$$\frac{T}{T_1} = \frac{m}{m_1} = r,$$

where r *is* an integer. Thus, the estimation stride m_1 is restricted to divisors of m. This condition allows V and edf V to be calculated from the subsampled arrays w_{jm_1} and $R_w^d(jm_1)$, respectively. For each (N, m, m_1) combination, the number M of estimation summands to be used in (22) is calculated by (6), and the parameter p is defined by

$$p = \frac{M}{r} = \frac{M\tau_1}{\tau}. \tag{24}$$

A selection of edf values is shown in Table 2 for integral values of the power-law exponent β. Values were also computed for half-integral values of β, but are not shown; as expected, they interpolate the given values, For now, ignore the "%" rows, and observe how edf depends on r (or m_1) for $N = 1024$, m fixed. For each β, and for $m \geq 4$, it is clear that any value of r between 4 and m gives a value of edf that is nearly maximal for that m and β. As the listings for $m = 2$ and 3 show, we should take $r = m$ (i.e., $m_1 = 1$) in case $m < 4$. Here is an empirical result.

Assume an averaging time τ at most 1/4th the duration of the time-deviation record. For each discrete-time power law between white phase and random-walk frequency, any estimation period τ_1 between τ_0 and max $(\tau_0, T/4)$ that divides evenly into τ gives an MVAR estimator V whose edf is within 8 percent of the maximal value for r.

Table 2 shows that the variation of edf V with r is greatest for white phase ($\beta = 0$). Also, we see that p by itself is a rough estimate of edf V, especially for r in the recommended range $\min(4, m) \leq r \leq m$.

The choice of estimation period τ_1 may depend on a tradeoff between convenience and computational effort. For simplicity, one can always choose $\tau_1 = \tau_0$. If the data set is large, one can choose the largest acceptable value, $\tau_1 = T/4$, to minimize the number M of terms needed to calculate V from (7).

VI. Lower Bounds for MVAR EDF

The aim of this section is to uncover simple approximation formulas for edf V that can be used in practice in place of the exact summation (22). There are two rigorous lower-bound formulas that can serve this purpose.

A. Discrete Time

Up to now, we have concentrated on a time-domain formulation of edf V. The following result is proved by a frequency-domain argument, which is not given here.

Theorem 1 *Let the time residuals x_n be a discrete-time power-law process (sample period τ_0) with spectrum (11), where $-9/2 < \beta \le 0$. Assume that x_n has stationary Gaussian mean-zero second increments. Let $m = m_1 r$, where m_1 and r are positive integers. Using (4), (5), (7), and any positive integer M, form the MVAR estimator V with averaging time $\tau = m\tau_0$ and estimation period $\tau_1 = m_1\tau_0$. Then*

$$\operatorname{edf} V \ge \frac{M}{r}\frac{2I^2}{J}, \tag{25}$$

where

$$I = \int_0^{m/2} \frac{\sin^6(\pi x)}{[m \sin(\pi x/m)]^{2-\beta}}\, dx,$$

$$J = \int_0^{r/2} \frac{\sin^{12}(\pi x)}{[r \sin(\pi x/r)]^{4-2\beta}}\, dx.$$

In other words, we have a bound of form edf $V \ge \alpha p$, where $p = M/r$ as above. Tables of α vs β, m, and r can be generated by numerical integration.

B. Continuous Time

It is much easier to derive a useful lower bound for edf V^c. Let $p \ge 2$. From (23) we have

$$\begin{aligned}
\frac{1}{\operatorname{edf} V^c} &= \frac{2}{p}\left[\int_0^p \left(\rho_z^c(x)\right)^2 dx - \frac{1}{p}\int_0^p x\left(\rho_z^c(x)\right)^2 dx\right] \\
&\le \frac{2}{p}\left[\int_0^\infty \left(\rho_z^c(x)\right)^2 dx - \frac{1}{p}\int_0^2 x\left(\rho_z^c(x)\right)^2 dx\right].
\end{aligned}$$

This gives a bound of form

$$\operatorname{edf} V^c \ge \frac{a_0 p}{1 - \frac{a_1}{p}}, \quad p \ge 2. \tag{26}$$

The constants *a.* and al, which depend only on β, are computed by numerical integration. To use this expression as an approximation to edf *V, we* again let p== M/r.

C. An edf Approximation Strategy

The right sides of *(25)* and *(26)* can be regarded as candidate approximations for edf *V.* To assess their quality and to choose between them, tables were generated for a selection of N, 7n, and r. The following empirical strategy and error statement emerged.

Assume discrcle-time power-law phase noise with *exponent between **white phase and random-walk** frequency, **at least** 16 time-residual **points, an** averaging time τ at **most*** 1/5th ***the duration of the measurement, and an estimation period** τ_1 **between** τ_0 **and*** max $(\tau_0, \tau/4)$ ***that divides evenly into** τ.* In our ***notation,*** $0 \geq \beta \geq -4$, $N \geq 16$, $m \leq N/5$, and $m = rm_1$, *where **r is an integer between*** $\min(m, 4)$ ***and*** In.

*For **In** = 1 **or** 2, **the discrete-time lower bound (25) is used as an approximation for*** edf ***V.** In all other cases, **the continuous-time lower bound*** (26) ***is used.** The **relative** error of **this strategy is observed to be at most*** ± 11.1 *percent.*

*To **implement** this approximation in practice, use **the formula***

$$\text{edf } V \approx \frac{a_0 p}{1 - \frac{a_1}{p}}, \tag{27}$$

*where **p = M/r, M*** is ***obtained from (6), and** the coefficients* a_0, a_1, ***as junctions of*** m ***and*** β, *are **drown from** Table **5'.***

To balance the errors, it was found expedient to reduce the continuous-time edf approximation, for white phase only, by 5 percent. Table 3 includes this adjustment. Each "%" row in Table *2* shows the percentage errors of (27) for the row above. Table 3 represents the full range of observed errors.

VII. Compound Noise Spectra

The foregoing results and methods assume a discrete-time phase noise spectrum proportional to (11) for some fixed exponent β. If that were indeed the case, our statistical efforts ought to be directed toward estimating the two-parameter set consisting of β and the constant of proportionality. Instead, as usual, we find ourselves using parametric tools to evaluate the confidence of a nonparametric statistic. The value of edf V depends on β. What can we do in the presence of a polynomial phase noise model

$$S_x(f) = \sum_{\beta} g_\beta \left|\sin(2\pi f \tau_0)\right|^\beta , \qquad (28)$$

a finite sum of power-law spectra? Some help is given by the following theorem, which, although weak and perhaps obvious, is better than no knowledge at all about the situation.

Theorem 2 *Let the phase noise be a finite sum of independent component noises with stationary Gaussian mean-mm second increments. Form an* MVAR *estimator V from the given phase noise, and corresponding estimators V_k from the components. Then*

$$\operatorname{edf} V \geq \min_k \operatorname{edf} V_k .$$

In other words, we never do worse than the worst component.

To apply this theorem to the situation (28), assume that the component β values are all in some subinterval of [-4, 0] (the whole range, perhaps). Use (27) and Table 3 to compute edf V_β for each tabulated β in the subinterval, and take the smallest value as a *conservative* estimate of edf V. For example, if one believes that the noise has components between white phase and flicker phase, perhaps from prior knowledge, perhaps as evidenced by a log-log plot of mod u vs τ with slopes between -3/2 and -1, then one can minimize (27) over the first three rows of Table 3.

The proof of Theorem 2, although not difficult, is not given here. It can be generalized to AVAR estimators and other situations involving averages of the square of a stationary Gaussian mean-zero process. Its usefulness for MVAR is enhanced by the relatively weak dependence of estimator edf on β, as can be seen in Table 2. An inspection of similar tables for fully overlapped AVAR estimators [6][12] shows a much sharper dependence on β, especially for large τ/τ_0. For MVAR, minimizing over a set. of β causes a smaller loss of accuracy in the computation of estimator edf, given imperfect knowledge of the phase spectrum.

VIII. CONCLUSIONS

Although the overall problem of estimating modified Allan variance MVAR may appear to be more difficult than the same problem for conventional Allan variance AVAR, theoretical and numerical results calculated here from the third-difference approach show that in some ways the sit uation is actually reversed. An attractive expression for the equivalent degrees of freedom (edf) of MVAR estimators in the presence of power-law phase noise was derived, and simple approximations constructed. Numerical computations of edf yielded a rationale for choosing the est i mat ion period or stride: we found empirically that the use of an estimation period up to one-fourth the averaging time does not appreciably degrade the confidence of the estimator below that of the fully overlapped estimator. Often, in fact, there is no degradation. The computations also revealed that the extra filtering inherent in MVAR causes the edf of an estimator to be less sensitive to the power-law exponent than the edf of a typical AVAR estimator. . Consequent y, MVAR error bars can be more robust against spectrum uncertainties than AVAR error bars.

The most important limitation on these results, especially for long tests of oscillators, is that linear frequency drift must be negligible. If a drift rate is known from considerations external to the immediate data set, then one can remove it from the phase data, and we are back to the case of zero drift. For AVAR, it is known that estimation of drift from the data themselves, and removal therefrom, causes negative AVAR estimator biases that worsen as averaging time τ increases. The use of three-point [14] [15] or four-point [4] drift estimators, which extract a quadratic component of the t ime-residual sequence x_n, simplifies calculations of the mean and variance of estimators of AVAR with drift removed. I have no doubt that similar calculations for MVAR estimators can be made on the basis of four-point drift estimators that extract a *cubic* component of the sequence w_n of cumulative sums of x_n.

The work described in this study was performed by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

REFERENCES

[1] D. W. Allan, "Time and frequency (time-domain) characterization, estimation, and prediction of precision clocks and oscillators", *IEEE Trans Ultrasonics Ferroelectrics Freq Control*, vol UFFC-34, pp 647-654, 1987

[2] D. W. Allan, "Time and frequency metrology: current status and future considerations", *Proc 5th European Freq and Time Forum*, pp 1-9, 1991

[3] L. G. Bernier, "Theoretical analysis of the modified Allan variance", *Proc 41st Ann Freq Control Symp*, pp 116-121, 1987

[4] C. A. Greenhall, "The fundamental structure function of oscillator noise models", *Proc 14th Precise Time and Time Interval Applications and Planning Meeting,* pp 281–294, 1982

[5] C. A. Greenhall, "A structure function representation theorem with applications to frequency stability estimation", *IEEE Trans Instrum Meas,* vol IM-32, pp 364–370, 1983

[6] C. A. Greenhall, "Recipes for degrees of freedom of frequency stability estimators", *IEEE Trans Instrum Meas,* vol 40, pp 994-999, 1991

[7] C. A. Greenhall, "A shortcut for estimating the modified Allan variance", *Proc* 1992 *IEEE Freq Control Symp,* pp 262-264, 1992

[8] D. A. Howe, D. W. Allan, J. A. Barnes, "Properties of signal sources and measurement methods", *Proc 35th Ann Freq Control Symp,* pp 1-47, 1981

[9] N. J, Kasdin, T. Walter, "Discrete simulation of power law noise", *Proc 1992 IEEE Freq Control Symp,* pp 274-283, 1992

[10] P. Lesage, C. Audoin, "Characterization and measurement of time and frequency stability", *Radio* Science, vol 14, pp 521-539, 1979

[11] J. Rutman, F. L, Walls, "Characterization of frequency stability in precision frequency sources", *Proc IEEE,* VOI 79, pp 952-960, 1991

[12] S. R. Stein, "Frequency and time-their measurement and characterization", *Precision Frequency Control,* vol 2, E. A. Gerber and A. Ballato, eds, pp 191–232, 399–416, New York, 1985

[13] T. Walter, "Characterizing frequency stability: a continuous power-law model with discrete sampling", *IEEE Trans Instrum Meas*, vol **43, pp 69-79, 1994**

[14] M. A. Weiss, D. W. Allan, D. A. Howe, "Confidence on the second difference estimation of frequency drift", *Proc* 1.992 *IEEE Freq Control Symp*, pp 300–305, ***1992***

[15] M. A. Weiss, C. Hackman, "Confidence on the three-point estimator of frequency drift", ***Proc 24th Precise Time and*** Time ***Interval Applications and Planning Meeting***, pp ***451-460, 1992***

LIST OF TABLES

The Third-Difference Approach to Modified Allan Variance
Charles A Greenhall

$$Q_0 = 0,\ Q_n = \sum_{j=1}^{|n|} \frac{1}{2j-1}$$

noise	β	$R_w^{\mathrm{d}}(n)$	$R_w^{\mathrm{c}}(t)$
wh ph	0	$\frac{-\lvert n\rvert}{2\tau_0}$	$-\frac{\lvert t\rvert}{2}$
fl ph	-1	$\frac{(4n^2-1)Q_n}{4\pi\tau_0}$	$\frac{t^2 \log\lvert t\rvert}{2\pi}$
wh fr	-2	$\frac{\lvert n\rvert(n^2-1)}{12\tau_0}$	$\frac{\lvert t\rvert^3}{12}$
fl fr	-3	$\frac{-(4n^2-1)(4n^2-9)Q_n}{192\pi\tau_0}$	$\frac{t^4 \log\lvert t\rvert}{24\pi}$
rw fr	-4	$\frac{-\lvert n\rvert(n^2-1)(n^2-4)}{240\tau_0}$	$\frac{-\lvert t\rvert^5}{240}$
	nonintegral	$\frac{-\Gamma(1-\beta/2+n)}{2\tau_0\cos(\pi\beta/2)\Gamma(2-\beta)\Gamma(\beta/2+n)}$	$\frac{-\lvert t\rvert^{1-\beta}}{2\cos(\pi\beta/2)\Gamma(2-\beta)}$

Table 1: Generalized Autocovariances Derived from Discrete-Time and Continuous-Time Power-Law Phase Noises.

m	r	ml	M	P	beta 0.0	-1.0	-2.0	-3.0	-4.0	
		N = 1024								
1	1	1	1022	1022.	525.9	589.3	681.6	828.6	1022.	
					+0.0	+0.0	+0.0	+0.0	+0.0	%
2	1	2	510	510.0	262.6	31.0.1	380.8	459.1	432.3	
	2	1	1019	509.5	477.0	496.5	515.2	523.6	441.4	
					-0.1	-0.1	-0.1	-0.1	-0.1	%
3	1	3	339	339.0	174.6	210.3	260.1	304.4	271.0	
	3	1	1016	338.7	373.9	349.9	341.5	334.6	274.0	
					+11.1	-2.8	-3.9	-4.1	-5.0	%
16	1	16	62	62.00	32.15	39.5″?	48.69	55.29	47.55	
	2	8	123	61.50	58.06	59.26	59.68	58.73	47.60	
	4	4	245	61.25	72.74	61.99	59.93	58.57	47.43	
					+4.1	+0.1	-0.2	-0.3	-0.2	%
	8	2	489	61.13	77.60	62.26	59.84	58.46	47.33	
					-2.6	-0.6	-0.2	-0.3	-0.2	%
	16	1	977	61.06	78.88	62.26	59.78	58.40	47.29	
					-4.3	-0.″7	-0.2	-0.3	-0.2	%
128	1	128	6	6.000	3.375	4.061	4.909	5.552	4.766	
	2	64	11	5.500	5.754	5.841	5.857	5.716	4.535	
	4	32	21	5.250	7.005	5.922	5.706	5.525	4.367	
					+3.4	+0.4	-0.1	-2.3	+0.2	%
	8	16	41	5.125	7.354	5.840	5.599	5.417	4.277	
					-3.6	-0.3	-0.3	-2.5	+0.0	%
	16	8	81	5.063	7.410	5.784	5.542	5.361	4.231	
					-5.3	-0.4	-0.4	-2.6	+0.0	%
	32	4	161	5.031	7.405	5.755	5.513	5.332	4.207	
					-5.8	-0.4	-0.4	-2.6	+0.0	%
	64	2	321	5.016	7.394	5.739	5.498	5.318	4.196	
					-5.9	-0.4	-0.4	-2.6	+0.0	%
	128	1	641	5.008	7.386	5.732	5.491	5.311	4.190	
					-5.9	-0.4	-0.4	-2.6	+0.0	%
		N = 16								
1	1	1	14	14.00	7.475	8.327	9.561	11.51	14.00	
					-3.7	-3.1	-2.4	-1.4	+0.0	%
2	2	1	11	5.500	5.754	5.946	6.117	6.146	5.061	
					-10.6	-10.0	-9.2	-8.1	-5.9	%
3	3	1	8	2.667	3.815	3.526	3.386	3.224	2.508	
					+9.9	-2.0	-3.0	-7.2	-3.5	%
					wh ph	fl ph	wh fr	fl fr	rw fr	
							noise type			

Table 2: Values and Approximation Errors for MVAR Estimator edf

noise type	beta	m = 1 a0	a1	m = 2 a0	a1	m >2 a0	a1
wh ph	0.0	.51429	0	.93506	0	1.2245	.58929
	-0.5	.54277		.95407		1..0739	.59605
fl ph	-1.0	.57640		.97339		3..0030	.60163
	-1.5	.61688		.99246		.97732	.59769
wh fr	-2.0	.66667		1.0101		.96774	.57124
	-2.5	.72948		1.0237		.96102	.50974
fl fr	-3.0	.81057		1.0266		.94663	.41643
	-3.5	.91389		.99981		.90604	.34276
rw fr	-4.0	1.0000		.86580		.76791	.411.15

Table 3: Coefficients for Approximating MVAR Estimator edf